A Forecaster Evaluation of Turbulence Algorithms: A Summary of the Winter 2001 Study

Jennifer Luppens Mahoney[1], Jamie Braid[2], Barbara Brown[2], Tressa Fowler[2], and Jamie Wolff[2]

October 2, 2001

1. Introduction

Over the past several years, the Quality Assessment Group (QAG), funded by the Federal Aviation Administration (FAA) Aviation Weather Research Project (AWRP) has been evaluating model-based forecasts of clear-air turbulence (CAT) both subjectively and objectively (Brown et al. 1999 and 2000 and Mahoney et al. 2001). The subjective assessment described here (hereafter Winter2001), second in the series, covers the period from 16 February - 22 April 2001. The first assessment (Mahoney and Brown 2000), conducted during the winter of 2000 (Winter2000), included evaluations from the forecasters at the Aviation Weather Center (AWC). The Winter2001 assessment was modified to include forecasters from Delta airlines as well as the AWC.

The goals of the subjective assessments are to 1) supplement the objective assessments with a meteorological classification of the turbulence events, 2) identify the frequency of the meteorological factors leading to turbulence, 3) obtain a subjective evaluation of algorithm performance, and 4) compare the differences/similarities of the quality of the turbulence forecasts between the objective and subjective assessments and use the information to improve the Integrated Turbulence Forecasting Algorithm (ITFA) that is being developed by the AWRP Turbulence Product Development Team (PDT).

2. Description of Evaluation

2.1 Approach

Based on results collected for Winter2000, the methodology for evaluation the study developed for Winter2001 was slightly modified to 1) supplement the AWC evaluations with forecaster evaluations by Delta airline forecasters, to utilize their experience with turbulence that affects large commercial aircraft; 2) conceal the identity of the turbulence algorithms that were being evaluated, to limit forecaster biases; 3) reduce the number of algorithms to be evaluated from 14 to eight so that specific information regarding each

[1] Forecast Systems Laboratory, 325 Broadway R/FS5, Boulder, CO 80303; Mahoney@fsl.noaa.gov.
[2] National Center for Atmospheric Research, Research Application Program, P.O. Box 3000, Boulder, CO 80303.

one could be obtained; and 4) extend the questionnaire so that further detail regarding the meteorological features and algorithm performance could be obtained.

The Winter2001 questionnaire was enhanced from the instrument used in Winter2000 (Mahoney and Brown 2000). As noted above, further specific information was requested regarding underlying causes of turbulence and the performance of the algorithms. Development of these enhancements was accomplished through collaboration with a variety of groups, including Delta airlines, the AWC, and the AWRP's Turbulence Product Development Team (PDT). An important enhancement was the coding of the identities of the specific algorithms, which increased the statistical validity of the study.

Seven weather forecasters from Delta airlines and eight forecasters from the AWC were involved in evaluating the eight CAT algorithms from 16 February – 22 April 2001. The algorithms were applied to the RUC-2 (Rapid Update Cycle, Version 2) model (Benjamin et al. 1998), with model output obtained from the National Centers for Environmental Prediction NCO (NCEP Computer Operations). Model forecasts issued at 1200, 1500, 1800, and 2100 UTC, with lead times of 0, 3, 6, 9, and 12 hours were included in the evaluation. Algorithm displays were broken down into 5,000-ft layers between 18,000 and 42,000 ft.

Displays of the various CAT algorithms were created at NCAR and made available to the forecasters through a Web-based graphical user interface. The forecasters were asked to view the displays each day and compare the output from these model-based forecasts to their assessment of the location of CAT, its intensity, spatial and temporal location, and its source (e.g., mountain waves). Forecasters were allowed to use all available sources of data and observations [e.g., pilot reports (PIREPs), satellite data, model forecasts] to evaluate these CAT features. The questionnaire also requested information concerning characteristics of the CAT and the underlying weather situation. However, since the process was voluntary, only a subset of turbulence cases were classified. Also, since several forecasters were evaluating the weather and the algorithms, some cases were classified more than once. At the end of the evaluation, the questionnaires were returned to FSL and NCAR for analysis.

2.2 Tools

The tools used during the evaluation included a web-based interface that allowed the forecasters to view the output from the eight turbulence algorithms and a questionnaire asking the forecasters to address the meteorological aspects of the turbulence event and the performance of the algorithms. Each tool is discussed in this section.

2.2.1 Web tools

Web-based tools were developed allowing forecasters to view displays of the various CAT forecasts [http://www-ad.fsl.noaa.gov/afra/rtvs; link turbulence; link Exercise 2001]. Figure 1 shows an example of a turbulence forecast provided by the Ellrod index where the panel represents a specific flight level range from 22,000 – 26,000 ft. On the web site, for each algorithm, issue, and lead-time, a panel of nine displays was presented. By double-clicking one of the nine panels, the forecasters were able to obtain larger views of each panel allowing them quick assess to the output from each algorithm. On the displays, the algorithms were identified by code number only (i.e., not by name), in order to reduce the impacts of forecaster biases on the study. The forecasters were not informed regarding which code number was associated with each algorithm.

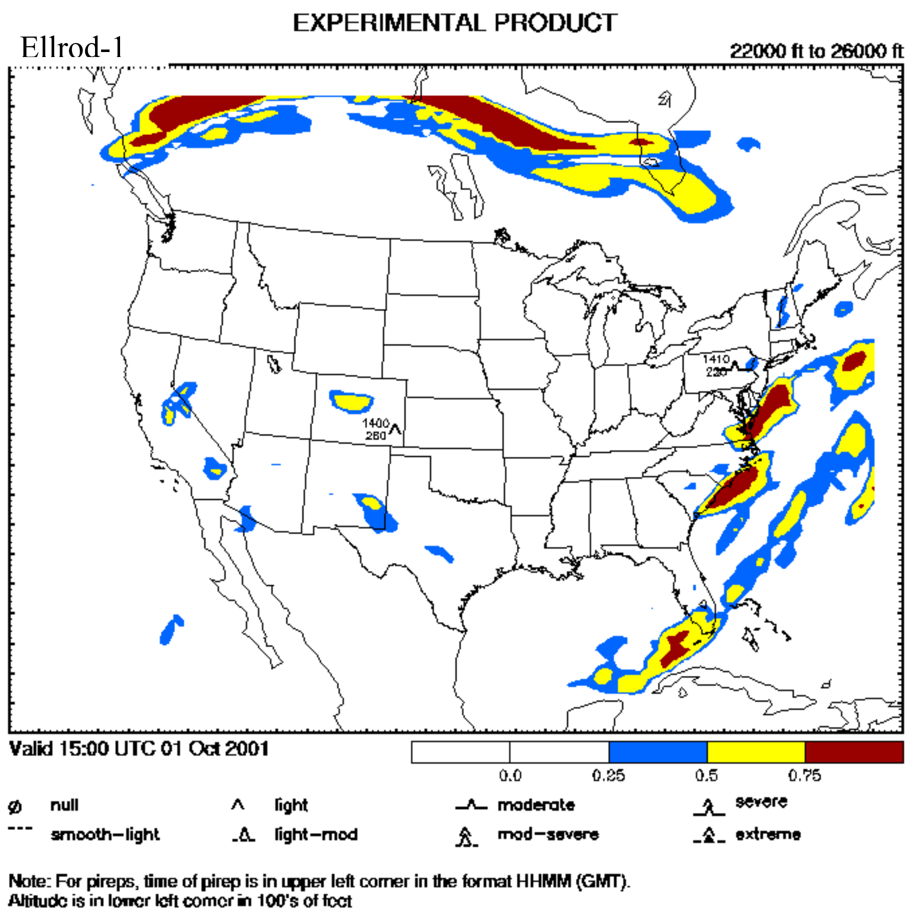


Figure 1. One of the nine display panels provided to the forecasters, where each panel represented a different flight level. This particular panel represents the experimental Ellrod index, 1200 UTC issue and 3-h lead for 1 October 2001. Color indicates turbulence intensity.

2.2.2 Questionnaire

The questionnaire addressed two main topics: weather classification, and algorithm assessment. In Section I of the questionnaire, forecasters addressed the severity, causes, location, and time of day the turbulence events occurred. In Section II of the questionnaire, forecasters evaluated whether the CAT forecasts captured the turbulence well, or did not capture it well. In addition, the respondents were asked to specifically indicate how well the algorithms performed in terms of spatial location and extent of predicted turbulence (vertically and horizontally), over- and under-forecasting, and severity. The questionnaires were created with guidance from other members of the Turbulence PDT, as well as Delta airlines and AWC staff, and then were modified based on feedback from the forecasters prior to the start of the evaluation. The questionnaire was an enhanced version of the questionnaire used for Winter2000, with greater detail requested regarding the underlying weather and algorithm performance. To be consistent with the algorithm displays, the algorithms were identified only using a code number. The questionnaires used for AWC and Delta are provided in Appendices A and B, respectively. Comments that were included in the AWC responses are summarized in Appendix C.

2.3 Description of turbulence algorithms

The eight algorithms considered in the subjective evaluation are described in this Section. Because the Upper-level Turbulence (Ulturb) forecasting index was not available to the forecasters throughout the entire evaluation period, results pertaining to that algorithm were excluded from the evaluation.

DTF3: The DTF3 ("Diagnostic Turbulence Formulation") algorithm was developed to take into account several sources of turbulent kinetic energy (TKE) in the atmosphere (e.g., upper fronts), with the output in terms of TKE (Marroquin 1995, 1998).

Ellrod-1: This index was derived from simplifications to the frontogenetic function. As such it depends mainly on the magnitudes of the potential temperature gradient, deformation and convergence (Ellrod and Knapp 1992).

Horizontal Shear (HS): This is the shear caused by the horizontal wind on a constant theta surface. It is largest near the horizontal boundaries of jet streams. Horizontal shear has been identified by Dutton (1980) as a major contributor to turbulence encounters by aircraft.

ITFA: The ITFA (Integrated Turbulence Detection and Forecasting Algorithm) forecasting technique uses fuzzy logic to integrate available turbulence observations (in the form of PIREP data) together with a suite of turbulence diagnostic algorithms (a superset of algorithms used in the verification exercise and others) to obtain the forecast (Sharman et al. 1999, 2000). This algorithm is under development by the Turbulence PDT.

Richardson Number (Rich): Theory and observations have shown that at least in some situations patches of CAT are produced by what is known as Kelvin-Helmholtz (KH) instabilities. This occurs when the Richardson Number, the ratio of the local static stability to the local shears, becomes small. Therefore, theoretically, regions with small Rich should be favored regions of turbulence (Drazin and Reid 1981; Dutton and Panofsky 1970; Kronebach 1964).

Temperature Gradient (TG): This is the horizontal gradient of temperature on a constant theta surface. It is a measure of deformation and also vertical wind shear from the thermal wind relation. This technique was recommended by forecasters at Delta Airlines as a good indicator of turbulence locations.

Ulturb: Ulturb, developed by Don McCann (1997), attempts to correlate unbalanced (i.e. nongeostrophic) flow to regions of clear-air turbulence. Three different measures of this imbalance are computed and the maximum of these relates to turbulence potential. The correlation between unbalanced flows and turbulence is supported at least qualitatively from numerous field experiments, both over the continental U.S. and the N. Pacific (Knox 1997).

Vertical Wind Shear (VWS): Helmholtz identified wind shear as a destabilizing force in the atmoshpere. This can be seen from its inverse relation to Richardson's number: large values favor a small Richardson's number, which in turn produces turbulence in stratified fluids (Drazin and Reid 1981; Dutton and Panofsky 1970).

3. Results

The results from the questionnaires from the Delta and AWC forecasters are summarized in this Section.

3.1 General comments

A total of 81 questionnaires were collected from the AWC forecasters from 16 February–31 March 2001; representing 39 days, with eight forecasters participating in the evaluation. Eighty-four questionnaires were collected from the Delta forecasters from from 26 February–22 April 200; representing 49 days, with eight forecasters participating in the evaluation. Since filling out a questionnaire was voluntary, only a subset of the total number of turbulence events was classified. In some cases, more than one forecaster completed a questionnaire for the same day.

3.2 Weather classification

The results presented in this Section pertain to the weather that is associated with turbulence, as considered in Part I of the questionnaire. The forecasters categorized the turbulence according to extent, severity, cause, location (both horizontally and vertically), time, duration, and relationship to a ridge/trough system. The manner in which the questionnaires were structured encouraged forecasters to describe turbulence that occurred over a period of a "day" rather than for a specific turbulence case. As a result every response, including those with multiple answers, were included in the tally and summarized in the diagrams provided in the following sections. In the future, the questionnaire will be structured so that information pertaining to individual weather events or cases can be identified.

3.2.1 Severity and causes of turbulence

The days with observed turbulence are categorized in Table 1 by extent (e.g., Big, Moderate, Small) and severity of turbulence. As a result of the subjectivity and daily classification of the turbulence, some of the days may appear in more than one category in the Table. In addition, AWC forecasters may have categorized a particular day differently than the Delta forecasters. This list is provided so that the developers of the algorithms can easily investigate cases that may be important to the development of ITFA. Of the days listed in Table 1, most had a turbulence severity of moderate, although, a large number of the days listed in Table 1 did severe reports of turbulence. A very small number of days during the evaluation were classified with having a severity of light or light to moderate. The light to moderate turbulence days were most often small in areal extent.

Table 1. A summary of days categorized by the areal extent of the turbulence activity and severity.

Delta		**AWC**	
Big Days			
Delta - Date	**Severity**	**AWC - Date**	**Severity**
March 12	Moderate	February 16	Severe
March 2	Severe	February 24	Severe
March 3	Severe	March 1	Severe
March 4	Severe	March 2	Severe
March 10	Severe	March 3	Severe
March 13	Severe	March 13	Severe
March 14	Severe	March 14	Severe
March 29	Severe	March 20	Severe
April 10	Severe	March 21	Severe
April 11	Mod - Severe		
Moderate Days			
Delta - Date	**Severity**	**AWC - Date**	**Severity**
March 13	Unknown	February 16	Moderate
April 12	Light	February 19	Moderate
February 27	Moderate	March 1	Moderate
February 28	Moderate	March 3	Moderate
March 6	Moderate	March 6	Moderate
March 7	Moderate	March 8	Moderate
March 9	Moderate	March 9	Moderate
March 10	Moderate	March 12	Moderate
March 11	Moderate	March 21	Moderate

March 15	Moderate	March 27	Moderate
March 21	Moderate	March 28	Moderate
March 26	Moderate	March 31	Moderate
March 27	Moderate	February 27	Severe
April 4	Moderate	February 28	Severe
April 7	Moderate	March 3	Severe
April 12	Moderate	March 7	Severe
April 22	Moderate	March 10	Severe
March 1	Severe	March 12	Severe
March 3	Severe	March 16	Severe
March 4	Severe	March 16	Mod-Severe
March 5	Severe		
March 6	Severe		
March 8	Severe		
March 9	Severe		
March 19	Severe		
March 22	Severe		
April 3	Severe		
April 5	Severe		
April 6	Severe		
April 7	Severe		
April 10	Severe		
March 5	Mod - Severe		
March 18	Mod - Severe		
March 28	Mod - Severe		
Small Days			
Delta - Date	**Severity**	**Delta - Date**	**Severity**

March 31	Light	March 9, 2001	Unknown
February 26	Moderate	February 16	Light
February 27	Moderate	February 17	Moderate
March 7	Moderate	February 18	Moderate
March 8	Moderate	February 21	Moderate
March 15	Moderate	February 22	Moderate
March 16	Moderate	February 23	Moderate
March 17	Moderate	February 26	Moderate
March 19	Moderate	February 27	Moderate
March 22	Moderate	March 4	Moderate
March 23	Moderate	March 5	Moderate
March 25	Moderate	March 6	Moderate
March 31	Moderate	March 7	Moderate
April 1	Moderate	March 11	Moderate
April 2	Moderate	March 12	Moderate
April 4	Moderate	March 17	Moderate
April 8	Moderate	March 18	Moderate
April 9	Moderate	March 19	Moderate
April 13	Moderate	March 22	Moderate
April 16	Moderate	March 24	Moderate
March 6	Severe	March 25	Moderate
March 21	Severe	March 26	Moderate
March 24	Severe	March 28	Moderate
April 2	Severe	February 23	Severe
April 8	Severe	March 11	Severe
April 12	Severe	March 22	Severe
March 17	Mod - Severe	March 5	Light-moderate
March 20	Mod - Severe	March 25	Moderate
April 1	Mod - Severe		

The forecasters were asked to identify the meteorological features they believed were major causes of turbulence. The results are summarized by the pie charts displayed in Figs. 2 and 3. The Delta and AWC forecasters overwhelmingly identified the jet stream as the major cause of turbulence, accounting for nearly half of the responses (44% and 48% respectively). Although, both forecasting groups identified upper-level though as a main source of turbulence, the AWC forecasters identified this source more than twice as often as the Delta forecasters (30% compared to 18%). Convection and the upper-level ridges also contributed to the development of turbulence, but only 6 – 19% of the time. Delta and AWC forecasters agreed that mountain waves were an infrequent cause of turbulence and only occurred on 2 – 3 % of the days.

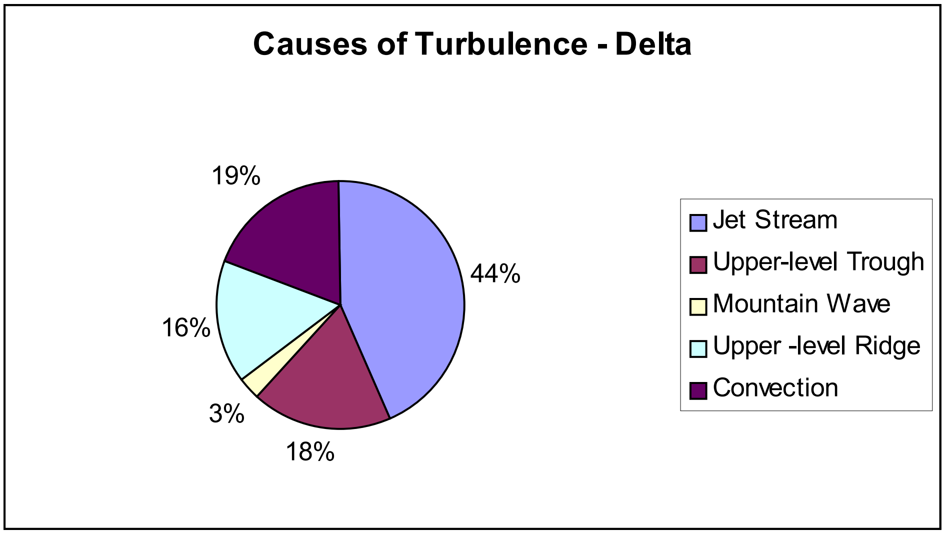


Figure 2. Pie chart summarizing the causes of turbulence as defined by the Delta forecasters. Large blue area is the jet stream (44%), red area located to the southwest is the upper-level trough (18%), purple area located to the northwest is convection (19%), light blue area represents the upper-level ridge (16%), and small neutral area is mountain wave (3%).

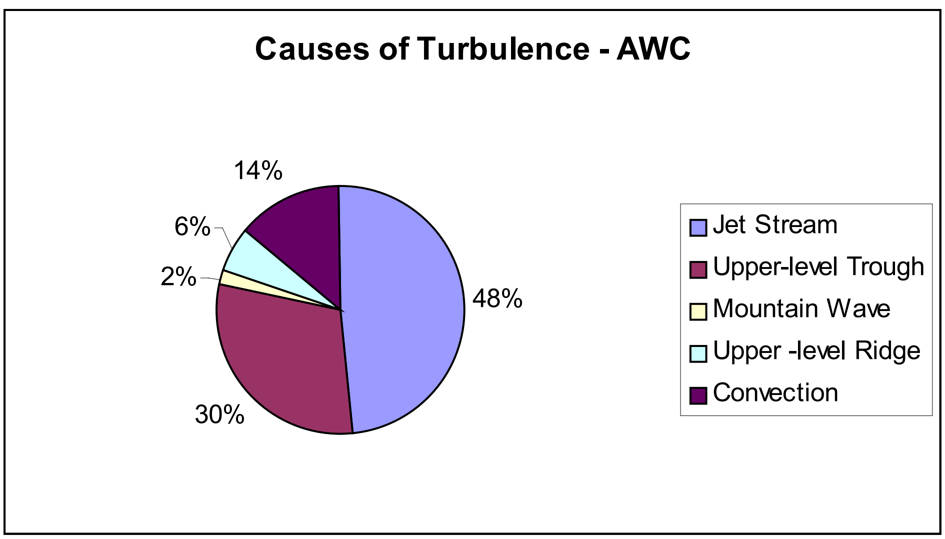


Figure 3. Same as Fig. 1, except for AWC responses.

3.2.2 Location and duration of turbulence

Cloud features are extremely important in the development of turbulence since the physical mechanisms that cause turbulence in clouds and in regions that are free of clouds are often very different. Therefore, the Delta and AWC forecasters were asked to categorize the turbulence with respect to location of the clouds. Four choices (i.e., *clear of clouds, near clouds, in clouds*, and *can't determine*) were provided to the forecasters in the questionnaire. The responses summarized in Fig. 4 include all single and combinations of responses.

The majority of responses from the Delta forecasters indicated that turbulence occurred equally in clear (38) as in cloudy regions (39). Although the forecasters indicated that on 30 days, the turbulence was located near the clouds. The AWC forecasters more often identified turbulence near clouds, but also recognized that many of the turbulence events on a particular day also occurred in clear regions (19 responses). Both groups agreed that the turbulence occurred most commonly near or in clouds more often, with 64 and 67 responses from the AWC and Delta forecasters respectively, than in clear regions. The days with turbulence categorized by location with respect to clouds are listed in Table 2. The dates of the turbulence days listed in Table 2 are the days when the forecasters provided only a single response to the question.

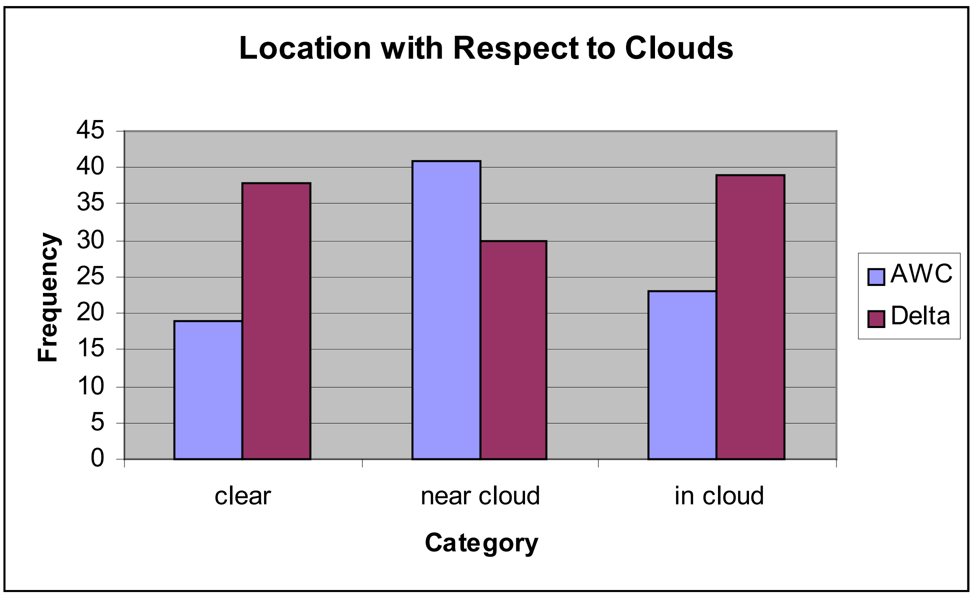


Figure 4. Bar chart showing the frequency of turbulence events with respect to the location of the clouds, as categorized by the Delta (red, right) and AWC (blue, left) forecasters.

In addition to categorizing the turbulence with respect to the location of clouds, forecasters were also asked to identify regions on a map where the turbulence was located. The choices provided on the questionnaire included: over the entire U.S., northern half of U.S., southern half of U.S., East, Central, and West. A map of the U.S. was also included on the questionnaire to provide an easy way for forecasters to represent

Table 2. A summary of turbulence days categorized by location with respect to clouds.

Delta		
Clear	**Near Cloud**	**In Cloud**
February 26 March 1 March 3 March 4 March 5 March 6 March 13 March 14 March 15 March 17 March 27 March 31 April 1 April 2 April 8 April 16	March 6 March 10 Mach 25 April 12 April 13	March 6 March 7 March 13 March 14 March 22 March 29 April 12 April 22
AWC		
Clear	**Near Cloud**	**In Cloud**
February 17 February 19 February 27 March 6 March 12 March 13 March 18 March 20 March 24 March 25 March 26 March 31	February 16 February 21 February 22 February 23 February 26 February 27 February 28 March 1 March 1 March 3 March 4 March 11 March 12 March 13 March 14 March 16 March 19 March 20 March 21 March 22 March 24 March 27 March 28	February 18 February 23 February 27 March 1 March 5 March 6 March 7 March 10 March 11 March 12 March 22 March 23 March 24

the location of the turbulence. These results were difficult to described and represent in this analysis. Therefore, we suggest that the reader refer to the raw questionnaires (available from the authors) for more detail. The location of the turbulence occurred most often over the Central region as categorized by the AWC forecasters and shown in Fig. 5. The Delta forecasters generally identified the turbulence over the entire U.S., but also frequently identified turbulence over the East, West, and Central regions of the U.S. Very few AWC or Delta forecasters classified location of the turbulence over the Northern or Southern half of the U.S., which suggests that these selections could possibly be removed from future questionnaires.

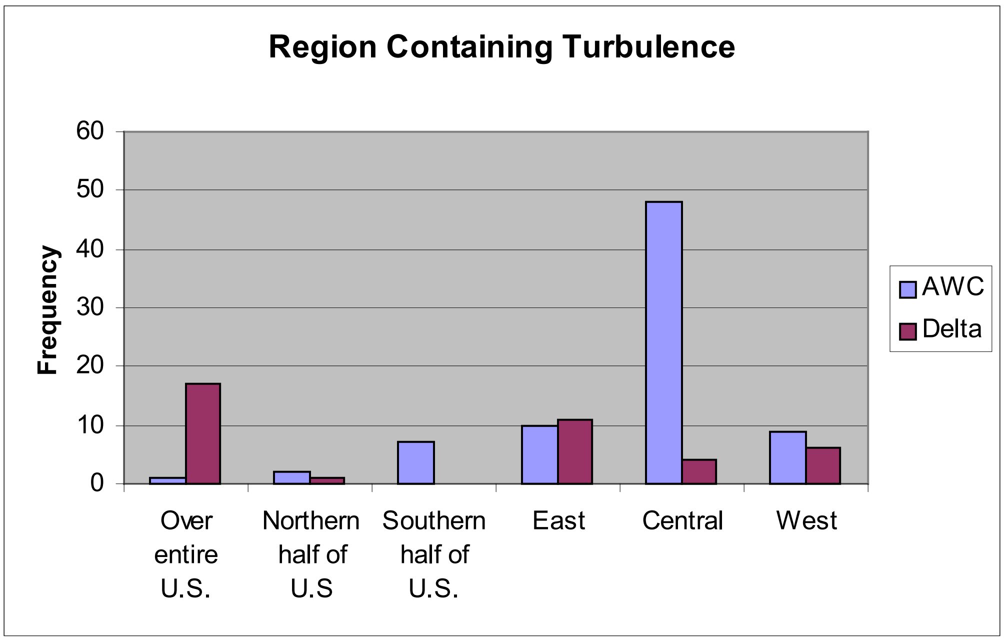


Figure 5. Bar chart showing the frequency of turbulence located in defined regions of the U.S., as categorized by Delta (red, right) and AWC (blue, left) forecasters.

As a supplement to questions regarding the location of turbulence by region, the forecasters at Delta and AWC were asked to categorize the observed turbulence by altitude. The responses are shown in Fig. 6. The most common altitude with turbulence was flight level 260-300 as defined by the Delta forecasters. Although AWC forecasters identified a large number of days with turbulence at this interval, the AWC forecasters found that the largest number of days had turbulence at altitudes that ranged from flight level 300-340. Overall, more than 80% of the days with turbulence events defined by either forecast group occurred between flight levels 200 and 380.

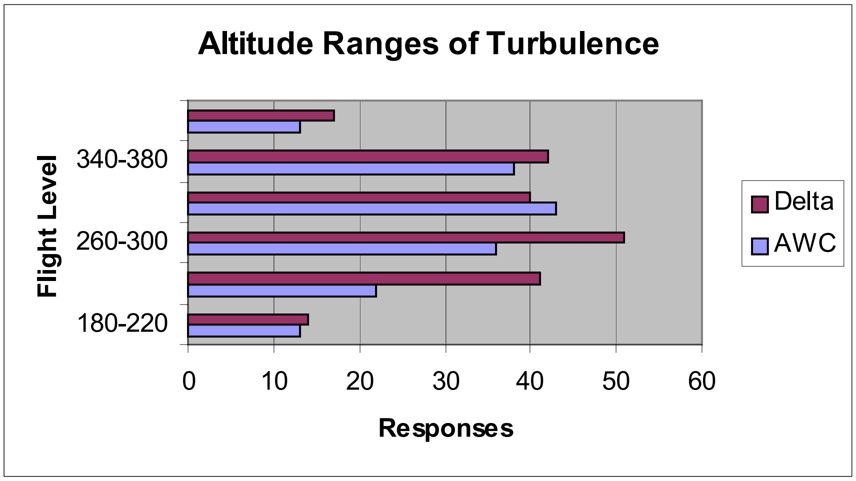


Figure 6. Bar chart of altitude ranges for observed turbulence as categorized by Delta (red, top) and AWC (blue, bottom) forecasters. Flight levels are listed along the y-axis in 5,000 ft bands and responses on the x-axis.

Figure 7 illustrates the time of day when turbulence was most often observed by the Delta and AWC forecasters. During the evaluation period, the Delta forecasters identified 1800–2100 UTC as the time of day with the majority of the turbulence events. However, their AWC counterparts observed half as many events during the same period. The peak time of day most often observed to contain turbulence as defined by AWC forecasters was 2100–0000 UTC; 6-h later than the period defined by Delta. However, both forecasting groups agreed that the majority of turbulence occurs between 1500 and 0000 UTC. The small number of responses identified between 0000–0300 UTC and 1200–1500 UTC may be in response to low numbers of PIREPs available during those hours, or to workload issues, rather than a reduction in the turbulence activity during those hours.

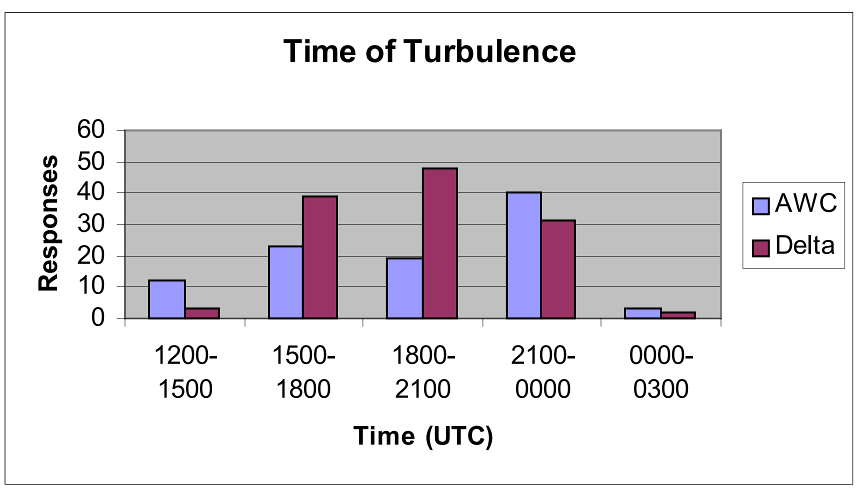


Figure 7. Bar chart of time of day when the major observed turbulence events occurred as categorized by Delta (red, right) and AWC (blue, left) forecasters. Times are in UTC.

3.2.5 Length of turbulence and location as related to trough/ridge pattern

Forecasters were asked to identify the length of time that the observed turbulence persisted. The results are summarized in Fig. 8. In general, most turbulence events persisted longer than 4 h, but often ended after a 12 h period. The Delta forecasters most often identified turbulence that persisted from 4 to 6 h, while the AWC forecasters identified turbulence events less then half as often for that category. Turbulence events defined by AWC forecasters tended to either continue for long periods which were generally beyond 6 h, or for short periods that were less than 3 h.

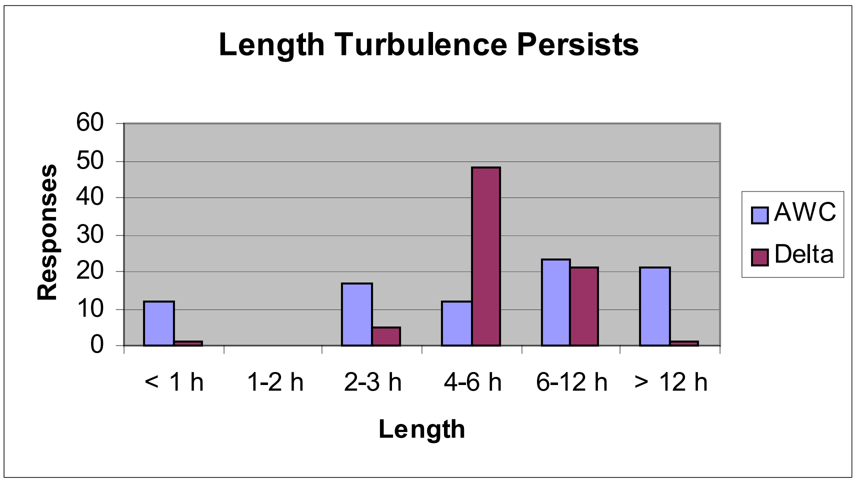


Figure 8. Bar chart of the length of time that the observed turbulence persisted as categorized by Delta (red, right) and AWC (blue, left) forecasters.

Finally, the Delta and AWC forecasters were also asked to identify the location of the observed turbulence with relation to a trough/ridge system, as defined by Fig. 9. The numbers in the diagram shown in Fig. 9 denote specific places within a trough/ridge system that may be conducive to turbulence. Specifically, numbers 1 and 5 signify the top or bottom of a ridge or trough, respectively. Numbers 3 and 6 symbolized NVA (negative vorticity advection) and PVA (positive vorticity advection) situations, and 2 and 4 indicated cutoff situations where 2 would be defined as a cutoff ridge and 4 a cutoff trough. The results of this analysis using this diagram are summarized in Fig. 10.

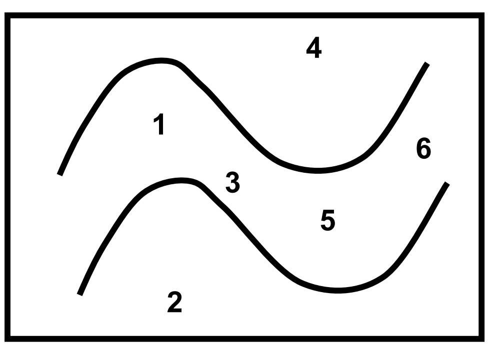


Figure 9. A graph used by the forecasters to identify the portion of a trough/ridge system where turbulence occurs most often. Numbers 1 and 5 are the top and bottom of the ridge/trough, respectively, 3 and 6 symbolize NVA and PVA, and 2 and 4 represent cutoff ridge/trough situations.

As shown in Fig. 10, the preferred located for the observed turbulence occurred consistently within the flow of the short wave ridge/trough system. Turbulence occurred most frequently at the top of the ridge, the bottom of the trough, and in NVA and PVA locations within the wave. Turbulence occurred less often in the cutoff ridge or cutoff trough locations, although some turbulence was identified in those regions.

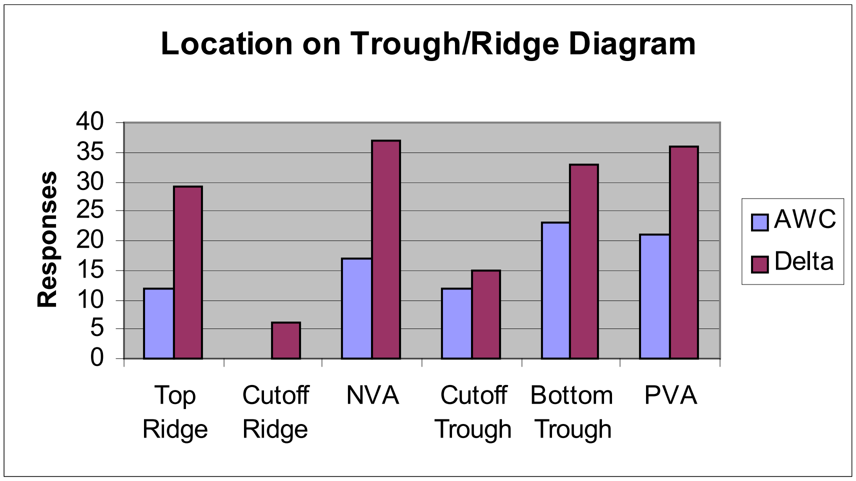


Figure 10. Bar chart describing the responses from the Delta (red, right) and AWC (blue, left) forecasters regarding the location of observed turbulence on the trough/ridge diagram (as shown in Fig. 9).

3.3 Algorithm assessment

The results presented in this section pertain to the quality of the turbulence algorithms, as considered in Part II of the questionnaire. The algorithms that were evaluated include: Ellrod-1, Richardson Number (Rich), DTF3, vertical shear (VWS), horizontal shear (HS), temperature gradient (TG), and ITFA. Ulturb, with only a few responses, was excluded from these analyses.

Figure 11 illustrates how well the forecasters believed the various algorithms captured turbulence. Percents of responses from Delta and AWC for each algorithm are displayed separately, with each bar totaling 100% of the responses from each forecasting group for each algorithm. Forecasters had four choices to describe algorithm performance: “captured turbulence well” (Well), “did not capture turbulence well” (Poorly), “forecast too much turbulence” (Overforecasted), and “provided a good indication of turbulence associated with mountain waves.” Responses in the mountain wave category were not included on the plot, since the number of responses in this category was extremely small.

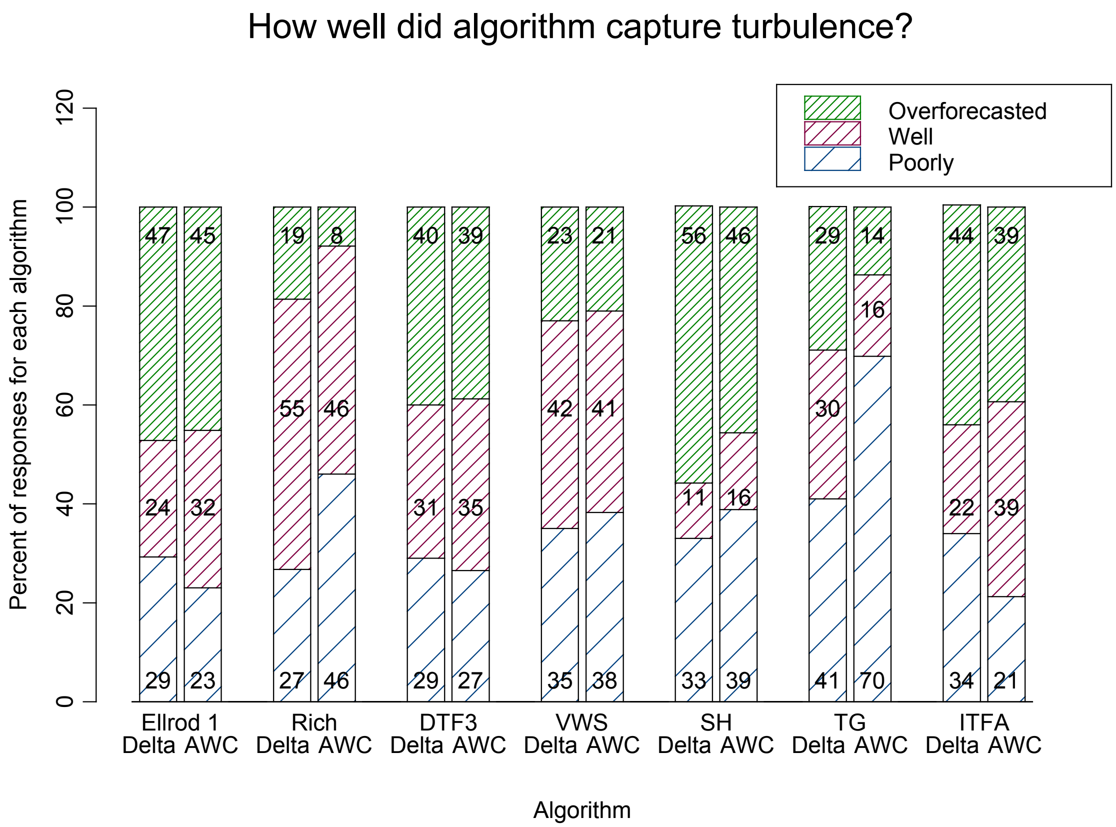


Figure 11. Stacked bar chart of algorithm performance for Ellrod-1, Rich, DTF3, VWS, HS, TG, and ITFA. Blue (wide) hatches indicate "poorly," red (medium) hatches indicate "well," and green (narrow) hatches indicate "overforecasted." Number of responses are represented by a percentage of the total responses. Delta bars are to the left and the AWC bars are to the right. Each set of 2 bars indicates a different algorithm.

Overall, the evaluations of each algorithm by Delta and AWC forecasters were fairly consistent. Rich was the algorithm judged least likely to overforecast turbulence, and most likely to capture turbulence well. HS was judged least likely to capture turbulence well and most likely to overforecast. TG was thought most likely to fail to capture turbulence.

The most noticeable differences between the Delta and AWC evaluations are for ITFA and TG. AWC forecasters believed that ITFA captured turbulence well nearly twice as frequently as Delta forecasters (39% vs. 22%). Conversely, Delta forecasters believed that TG captured turbulence well nearly twice as frequently as AWC forecasters (30% vs. 16%). While AWC and Delta forecasters agreed that TG was most likely to fail to capture turbulence well, the AWC forecasters judged that TG failed to capture turbulence well much more often (70% vs. 41%).

Forecasters were asked to assess the size of each algorithm's turbulence forecast. Frequencies associated with their assessments are displayed in Fig. 12. As in the previous figure, the percent of responses from Delta and AWC for each algorithm are displayed separately, with each bar totaling 100% of the responses from each location for each algorithm. Forecasters had three choices to describe the size of each algorithm's forecast area: "Too Small," "Correct size," and "Too Large."

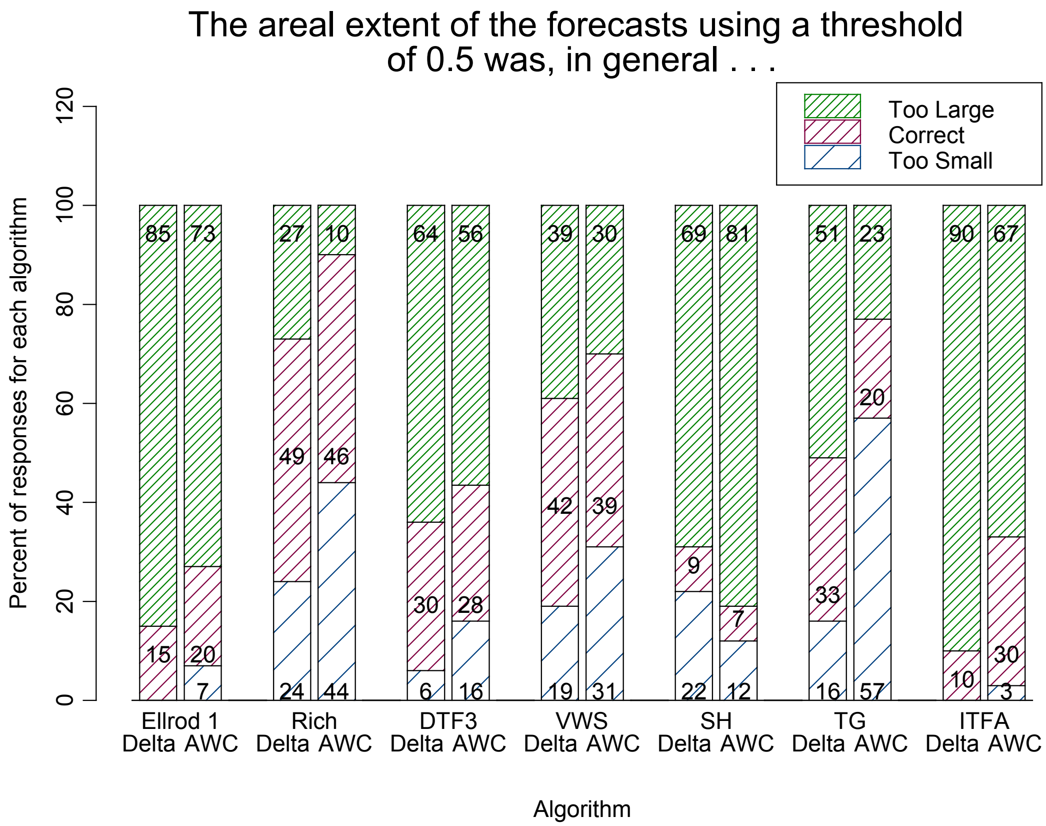


Figure 12. Same as Fig. 11, except for areal extent of turbulence.

Generally, the forecasters overwhelmingly believed the areal extent of the algorithms' turbulence forecasts to be too large. The only major exception is the evaluation of Rich by AWC forecasters. According to the AWC forecasters, Rich forecasts were too large only 10% of the time. However, they were too small nearly as often as they were the correct size (44% and 46%, respectively). Rich and VWS forecasts were judged to be of the correct size most often by both AWC and Delta forecasters. HS was least frequently of the correct size. TG and Rich were judged most likely to be too small while Ellrod-1 and ITFA were least likely to be too small.

Since most algorithms produce continuous turbulence forecast values, rather than yes/no forecasts, the threshold value applied can make a huge difference in determining the size/quality of the forecasts. For this analysis, thresholds from the algorithms producing similar values of PODy were used to evaluate the forecasts. However, as algorithm calibration may vary considerably, application of these thresholds to the algorithms does not ensure equitability in the analysis of forecast size.

In the analysis of altitude, shown in Fig. 13, Rich was once again favored by both the Delta and AWC forecasters. It was most often determined to have forecast turbulence at appropriate altitudes. The respondents also indicated that DTF3, Ellrod-1, and VWS forecasts were quite frequently located at appropriate altitudes. A large proportion of TG forecasts were deemed to be at altitudes that were too low. For the remaining algorithms, the incorrect forecasts were more often too high than too low.

Overall, the Delta forecasters were more likely than the AWC forecasters to judge that a forecast altitude was too high. In fact, for six of the seven algorithms (Ellrod-1, Rich, DTF3, HS, TG, ITFA), the Delta forecasters determined that the forecast altitude was too high more than twice as frequently as the AWC forecasters. Conversely, the AWC forecasters judged the forecasts to be too low at least twice as frequently as the Delta forecasters for all algorithms. However, the forecasters seemed to agree at least somewhat on how frequently the forecast altitude for each algorithm was appropriate.

As in the previous analysis, thresholds from the algorithms producing similar PODy values were used to make these comparisons. Again, due to calibration issues, application of the same threshold to all algorithms does not ensure equitability in the analysis.

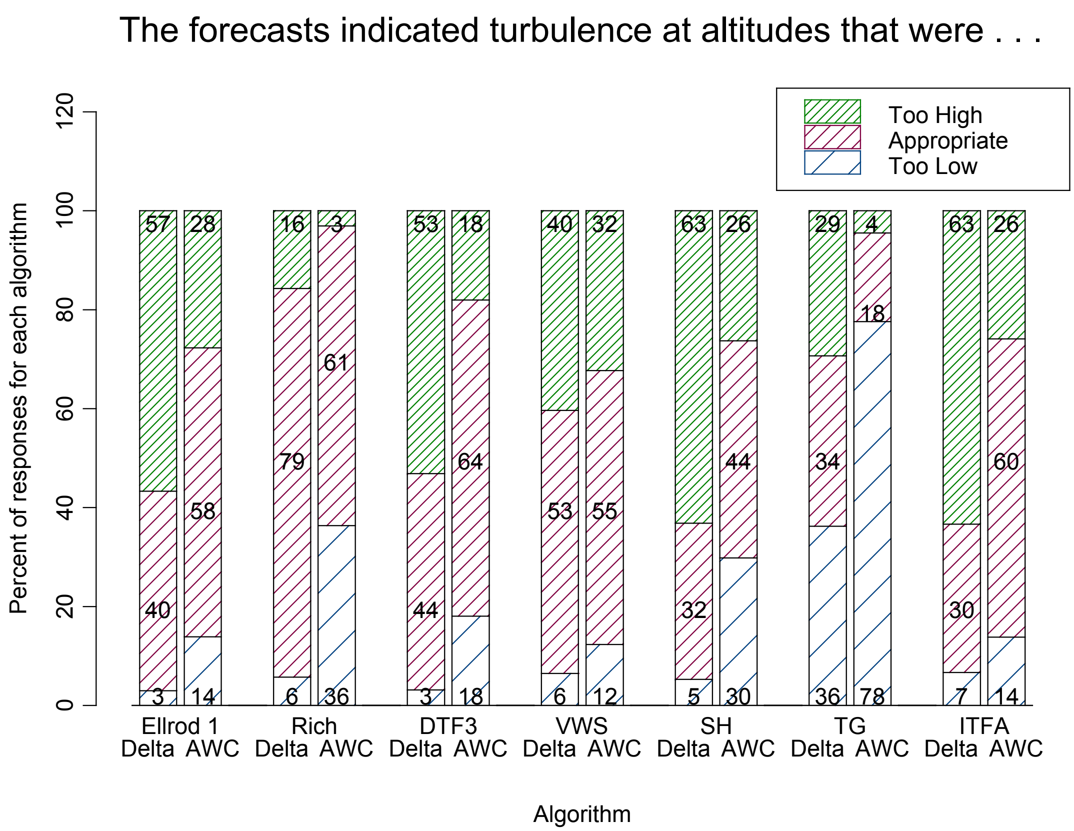


Figure 13. Same as Fig. 11, except for turbulence separated by altitude.

.

Generally, the respondents felt that the algorithms forecast turbulence severities that were too high, as shown in Fig. 14. Only Rich forecasted appropriate severities a large percentage of the time according to both AWC and Delta forecasters.

Delta forecasters indicated that all other algorithms forecast turbulence at severities that were too high in a huge proportion of cases. AWC forecasters believed that Ellrod-1, VWS and ITFA forecast appropriate severities a reasonable proportion of the time (49%, 56% and 46%, respectively). However, their counterparts at Delta did not appear to agree. The AWC forecasters believed that Rich, VWS, and TG forecast turbulence severities that were too low more often than the other algorithms. Delta forecasters indicated that low turbulence severities were rare for all algorithms.

The differences between the AWC and Delta forecasters' assessment of turbulence severity may have little to do with the algorithms. Perhaps it has more to do with the intended users of the forecasts. Delta forecasts are intended for professional pilots, usually flying large aircraft. AWC forecasts are intended for everyone, including pilots who have less flying experience and may be flying smaller aircraft.

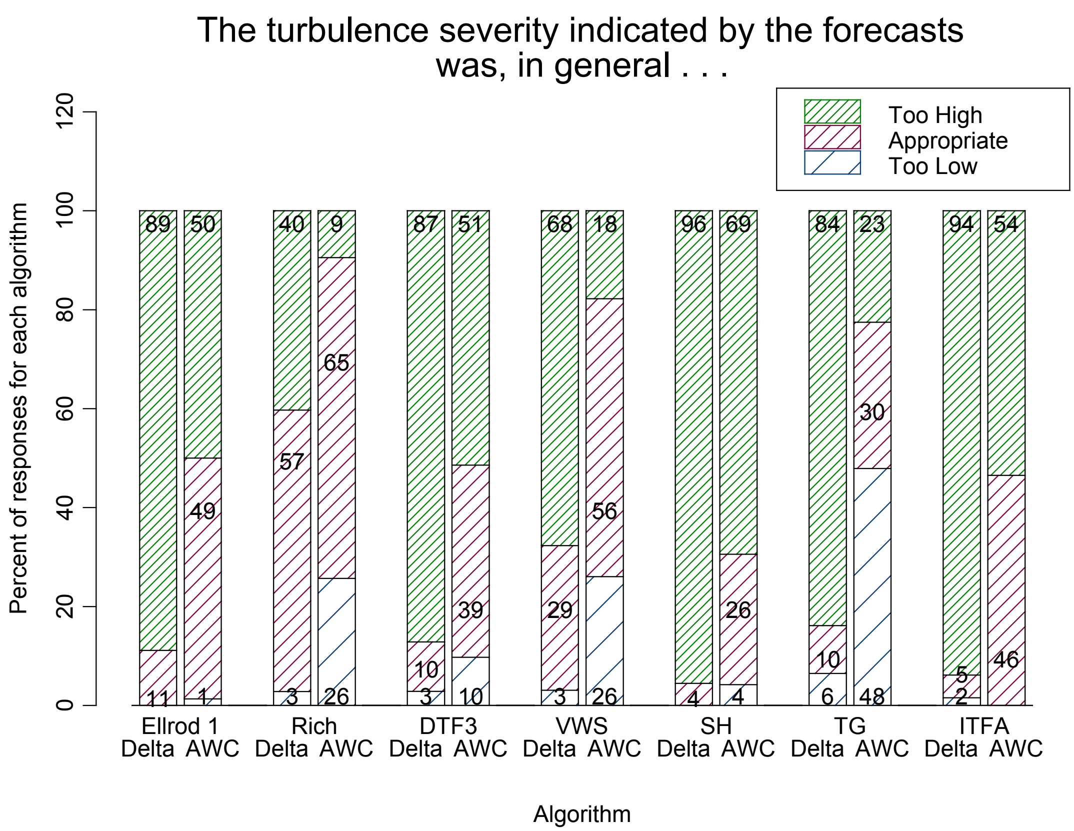


Figure 14. Same as Fig. 11, except for turbulence separated by altitude.

3.4 *Assessment of algorithms stratified by forecaster*

3.4.1 Delta forecasters

At Delta Airlines, seven forecasters participated in the survey. Figure 15 shows a pie chart for the percent of the surveys submitted by each Delta forecaster. Two forecasters, numbers 1 and 2 in Tables 3 and 4, account for over half (53%) of the surveys received from Delta. The remaining forecasters each completed only a small number of surveys. The two forecasters experienced somewhat different weather conditions during the times when they evaluated the algorithm performance. Not surprisingly, they also had different opinions about the performance of the algorithms.

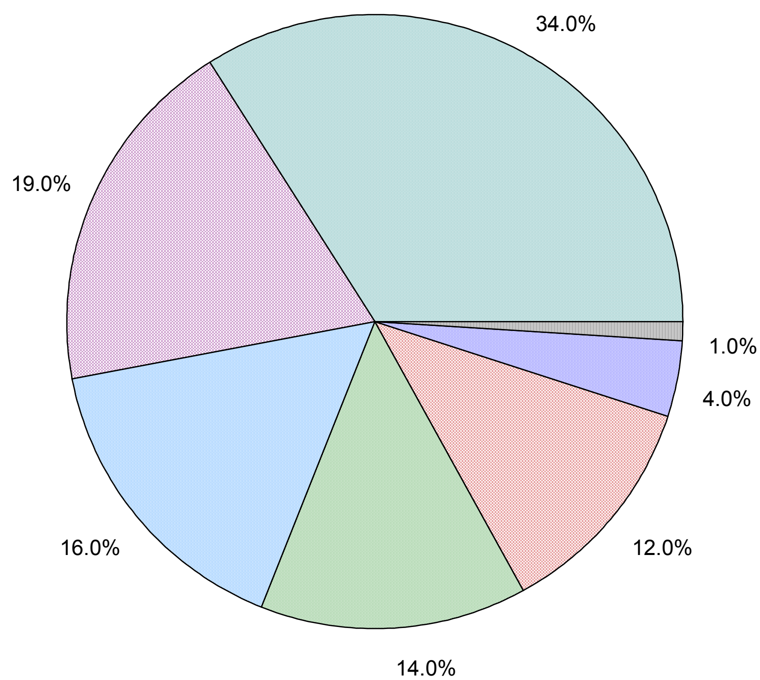


Figure 15. Percent of survey responses per forecaster - Delta Airlines.

With respect to the extent of the turbulence Table 3, Forecaster 1 (DF1) evaluated the algorithms on a higher (lower) percentage of "Moderate" ("Small") days than did Forecaster 2 (DF2). Both forecasters completed surveys on about the same proportion of "Big" Days (11% and 13%, respectively). During 60% of the cases when DF1 judged the algorithm performance, the maximum turbulence severity (Table 4) was Moderate to Severe or Severe. For DF2, turbulence of those severities occurred only 25% of the time.

Table 3. Forecaster -vs- Extent of Turbulence (Percent).

	Missing	Small Day	Small to Mod Day	Moderate Day	Mod to Big Day	Big Day	Total
Delta Forecaster #1	4%	43%	4%	39%	0%	11%	34%
Delta Forecaster #2	0%	56%	0%	25%	6%	13%	19%
Delta Forecaster #3	0%	31%	0%	62%	0%	8%	16%
Delta Forecaster #4	0%	30%	0%	40%	0%	30%	12%
Delta Forecaster #5	0%	33%	0%	67%	0%	0%	4%
Delta Forecaster #6	0%	33%	0%	42%	0%	25%	14%
Delta Forecaster #7	0%	0%	0%	100%	0%	0%	1%
Column Average	2%	39%	1%	42%	1%	14%	100%

Table 4. Forecaster -vs- Max Severity of Turbulence (Percent).

	Missing	Light	Moderate	Mod to Severe	Severe	Total
Delta Forecaster #1	4%	0%	36%	21%	39%	34%
Delta Forecaster #2	6%	0%	69%	6%	19%	19%
Delta Forecaster #3	0%	0%	15%	0%	85%	16%
Delta Forecaster #4	0%	10%	50%	0%	40%	12%
Delta Forecaster #5	0%	0%	67%	33%	0%	4%
Delta Forecaster #6	0%	8%	58%	0%	33%	14%
Delta Forecaster #7	0%	0%	100%	0%	0%	1%
Column Average	4%	2%	45%	10%	39%	100%

Figure 16 shows the percent of each forecasters' responses indicating how well the algorithm captures turbulence. Clearly, the forecasters only agreed on Rich. For all other algorithms, the forecasters gave very different responses regarding how well the algorithms captured turbulence. DF2 indicated that the algorithms captured turbulence well (overforecasted) a greater (lower) proportion of the time than DF1. Most likely, this analysis is confounded by the different weather conditions (extent and severity of turbulence) experienced by each forecaster.

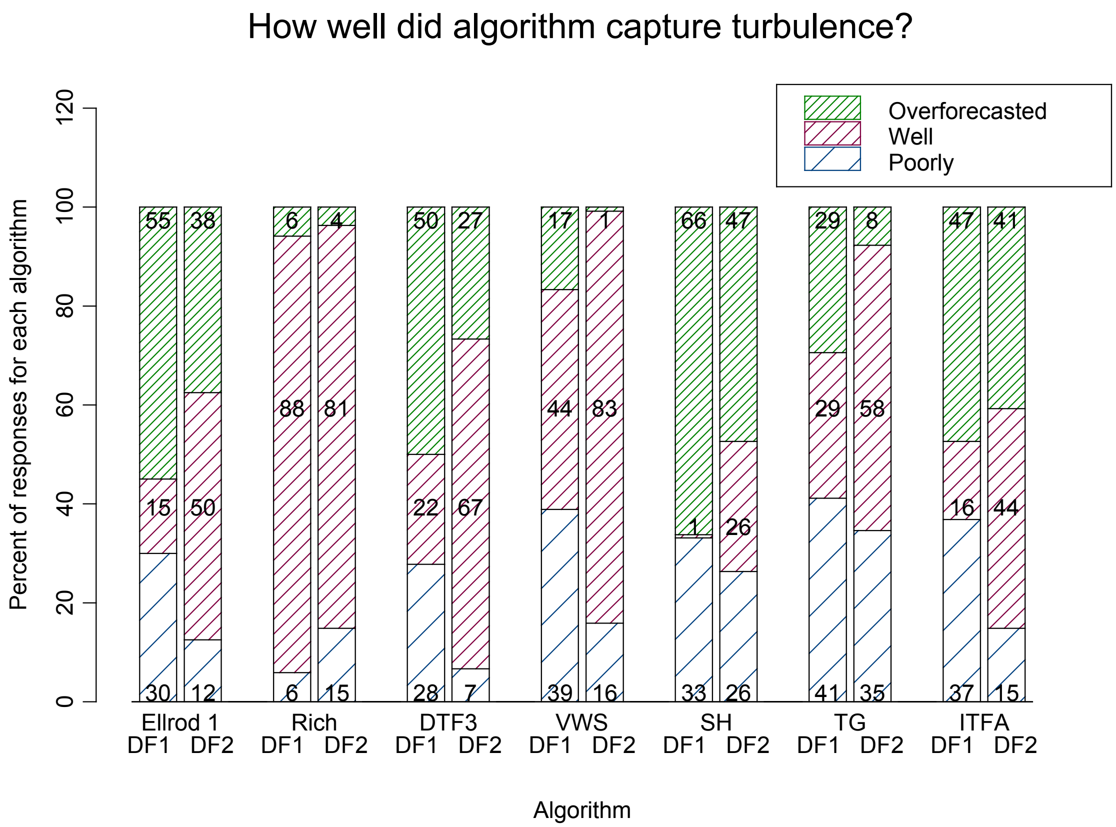


Figure 16. Same as Fig. 1, except for Delta forecasters.

3.4.2 AWC forecasters

At the AWC, nine forecasters participated in the survey. Figure 17 shows a pie chart for the percent of the surveys submitted by each AWC forecaster. As with Delta, two forecasters provided the bulk (43%) of the surveys. The two forecasters experienced somewhat different weather conditions during the times when they evaluated the algorithm performance. Not surprisingly, they also had different opinions about the performance of the algorithms.

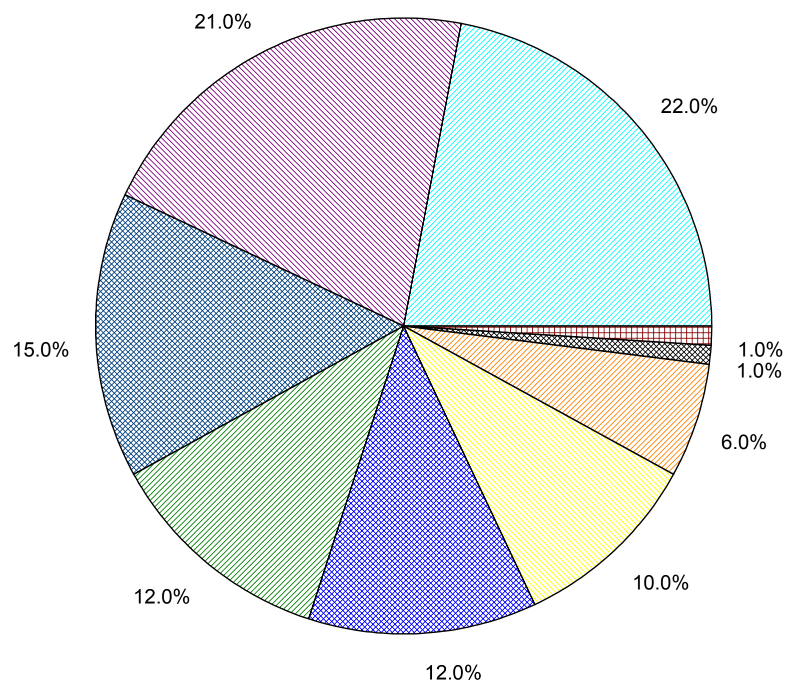


Figure 17. Percent of survey responses per forecaster - AWC.

With respect to the extent of the turbulence (Table 5), Forecaster 1 (AF1) evaluated the algorithms on a higher (lower) percentage of “Moderate” (“Small”) days than did Forecaster 2 (AF2). Both forecasters completed surveys on about the same proportion of “Big” Days (11% and 13%, respectively). During 60% of the cases when AF1 judged the algorithm performance, the maximum turbulence severity (Table 6) was Moderate to Severe or Severe. For AF2, turbulence of those severities occurred only 25% of the time.

Table 5. Forecaster -vs- Extent of Turbulence (Percent)

	Missing	Small Day	Moderate Day	Big Day	Total
AWC Forecaster #1	0%	44%	17%	39%	22%
AWC Forecaster #2	0%	53%	35%	12%	21%
AWC Forecaster #3	0%	33%	42%	25%	15%
AWC Forecaster #4	20%	40%	30%	10%	12%
AWC Forecaster #5	0%	0%	0%	100%	1%
AWC Forecaster #6	20%	70%	10%	0%	12%
AWC Forecaster #7	0%	12%	75%	13%	10%
AWC Forecaster #8	0%	0%	100%	0%	1%
AWC Forecaster #9	0%	40%	60%	0%	6%
Column Percent	5%	43%	34%	18%	100%

Table 6. Forecaster -vs- Maximum Turbulence Severity (Percent)

	Missing	Light	Light to Moderate	Moderate	Mod to Severe	Severe	Total
AWC Forecaster #1	0%	0%	0%	33%	0%	67%	22%
AWC Forecaster #2	6%	0%	6%	64%	0%	24%	21%
AWC Forecaster #3	0%	0%	0%	50%	0%	50%	15%
AWC Forecaster #4	20%	0%	0%	40%	0%	40%	12%
AWC Forecaster #5	0%	0%	0%	0%	0%	100%	1%
AWC Forecaster #6	10%	20%	0%	60%	0%	10%	12%
AWC Forecaster #7	0%	0%	0%	75%	12%	13%	10%
AWC Forecaster #8	0%	0%	0%	100%	0%	0%	1%
AWC Forecaster #9	0%	0%	0%	100%	0%	0%	6%
Column Percent	5%	3%	1%	55%	1%	35%	100%

Figure 18 shows the percentage of each forecasters' response indicating how well the algorithm captures turbulence. Each forecaster indicated that Rich captured turbulence well more often than any other algorithms. However, AF2 claimed this happened twice as

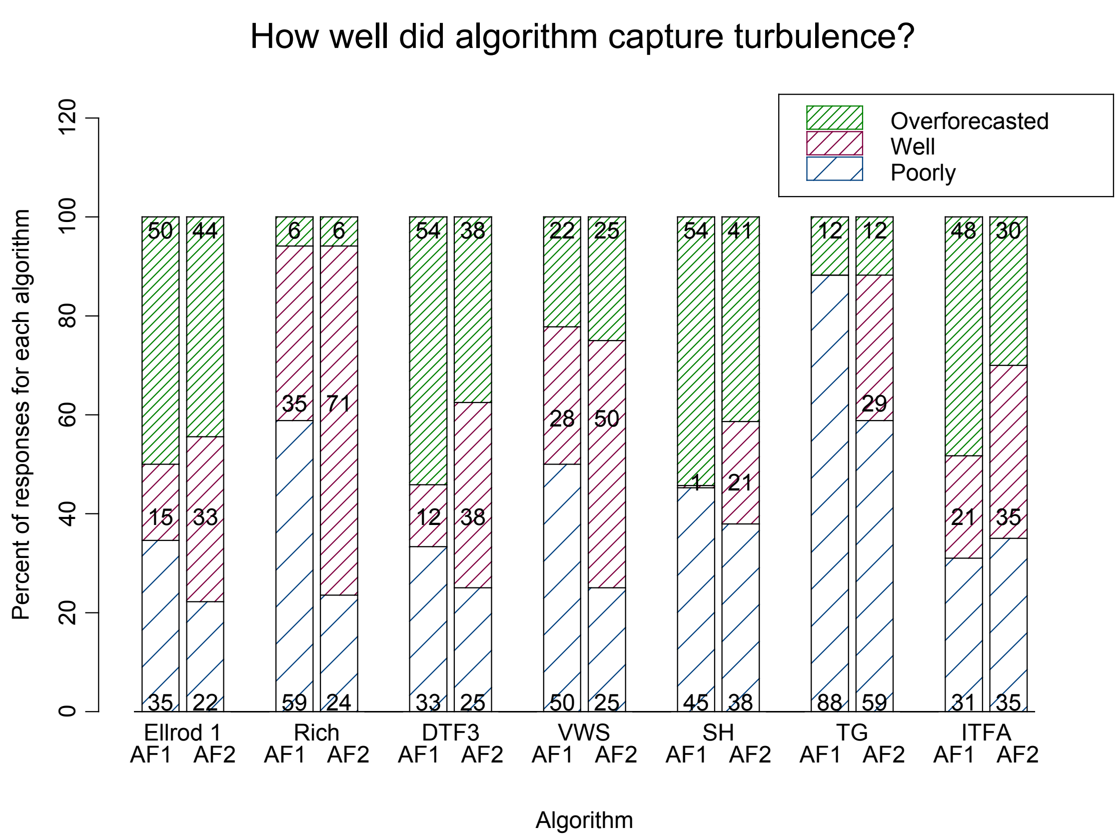


Figure 18. Same as Fig. 16, except for AWC forecasters.

frequently as AF1. The forecasters did not seem to agree in their assessments of any of the algorithms, although agreement is closest for ITFA. For all other algorithms, the forecasters gave very different responses regarding how well the algorithms captured turbulence. AF2 indicated that the algorithms captured turbulence well a greater proportion of the time than AF1. For all algorithms except ITFA, the proportion of forecasts that were classified as capturing turbulence poorly by AF1 was greater than the same classification by AF2. Again, this analysis is likely to be confounded by the different weather conditions (extent and severity of turbulence) experienced by each forecaster.

4. Summary

This report summarizes the results of the Winter2001 subjective evaluation of the eight CAT algorithms. This exercise was the second evaluation that has taken place over the past two winters. Forecasters from Delta airlines and the AWC participated in this evaluation during the period from 16 February – 22 April 2001.

The results provided a great deal of information regarding the important sources of turbulence and the performance of the algorithms. However, the structure of the questionnaire inhibited the ability to distinguish the location, severity, and time of individual turbulence events. Rather the responses captured the turbulence over a period of a day, often with several events occurring within that period.

Overall the results indicate that one of the top performers as indicated by AWC and Delta forecasters was the Richardson Number. The Richardson Number was valued positively in all areas of the evaluation. For instance, it was the algorithm believed to be least likely to over forecast turbulence, most likely to capture turbulence well, most often have the correct size of turbulence, and generally had turbulence at appropriate altitudes.

The quality of ITFA, our primary algorithm of interest, differed among the AWC and Delta forecasters. AWC forecasters felt that ITFA captured turbulence well nearly twice as often as the Delta forecasters. Both the AWC and Delta forecasters believed that the area of turbulence predicted by ITFA was too large and least likely to be to small. The AWC forecasters were more likely than the Delta forecasters to judge the forecast altitude and severity of ITFA as appropriate. These differences identified between the forecast groups may have to do with the intended users of the forecasts. Delta forecasts are intended for professional pilots usually flying large aircraft. AWC forecasts are intended for the general aviation community, which includes pilots with less flying experience and who may be flying smaller aircraft that are most sensitive to the less severe turbulence.

Some of the perceived differences among algorithms may also be due to the way the algorithms' forecasts were depicted (i.e., the thresholds selected). Future questionnaires should attempt to correct this problem.

Future work includes: allowing other airline groups to participate in the evaluation, modifying the questionnaire so that individual cases of turbulence can be identified, and begin comparing these results with those collected during the objective portion of the evaluation.

Acknowledgments

This research is in response to requirements and funding by the Federal Aviation Administration. The views expressed are those of the authors and do not necessarily represent the official policy and position of the U.S. Government.

We would like to thank the forecasters at the Aviation Weather Center and Delta Airlines who participated in this evaluation. Without them, this work could not have been performed. We also would like to thank Clinton Wallace at the AWC and Joe Luisi at Delta Airlines for being the focal points for this evaluation and for providing guidance to the forecasters when needed. We would especially like to express our appreciation to Jim Johnson of Delta Airlines for his help in motivating this study and the enhancements to the questionnaires. We also would like to thank the RTVS Team (Judy Henderson, Joan Hart, Andy Loughe, Beth Sigren, and Chris Fischer) who are the developers of the RTVS, Missy Petty at NCAR for enhancing and maintaining the algorithm displays, and Nita Fullerton (FSL) for her helpful review of this work.

References

Benjamin, S.G., J.M. Brown, K.J. Brundage, B.E. Schwartz, T.G. Smirnova, and T.L. Smith, 1998: The operational RUC-2. *Preprints, 16th Conference on Weather Analysis and Forecasting*, Phoenix, AZ, American Meteorological Society, 249-252.

Brown, B.G., J.L. Mahoney, R. Bullock, J. Henderson, and T.L. Kane, 1999: Turbulence algorithm intercomparison: 1998-1999 Initial Results. NOAA Technical Memorandum OAR FSL-25, Forecast Systems Laboratory, National Oceanic and Atmospheric Research, 63 pp.

Brown, B.G, J.L. Mahoney, J. Henderson, Tressa L. Kane, Randy Bullock, and Joan E. Hart, 2000: The turbulence algorithm intercomparison exercise: Statistical verification results. *Preprints, 9th Conference on Aviation, Range, and Aerospace Meteorology*, Orlando, Fl., American Meteorological Society.

Drazin, P.G. and W.H. Reid, 1981: **Hydrodynamic Stability**. Cambridge, 527 pp.

Dutton, M.J.O., 1980: Probability forecasts of clear-air turbulence based on numerical model output. *Meterol. Mag.,* **109**, 293-310.

Dutton, J. and H. A. Panofsky, 1970: Clear Air Turbulence: A mystery may be unfolding. *Science*, **167**, 937-944.

Ellrod, G.P. and D.I.Knapp, 1992: An objective clear-air turbulence forecasting technique: verification and operational use. *Wea. Forecasting*, **7**, 150-165.

Knox, J. A., 1997: Possible mechanism of clear-air turbulence in strongly anticyclonic flows. *Mon. Wea. Rev.,* **125**, 1251-1259.

Kronebach, G. W., 1964: An automated procedure for forecasting clear-air turbulence. *J. App. Met.,* **3,** 119-125.

Mahoney, J.L., and B.G. Brown, 2000: Forecaster Assessment of Turbulence Algorithms: A Summary of the Winter 2000 Study. NOAA Technical Memorandum OAR FSL-27, Forecast Systems Laboratory, National Oceanic and Atmospheric Administration, Department of Commerce, 134 pp.

Mahoney, J.L., B.G. Brown, R. Bullock, T.L. Fowler, C. Fischer, J. Henderson, and B. Sigren, 2001: Turbulence Algorithm Intercomparison: Winter 2001 Results. FAA Turbulence Product Development Team Report to the FAA Aviation Weather Research Program (Available from J. Mahoney, FSL, 325 Broadway, Boulder CO 80303).

Marroquin, A., 1995: An integrated algorithm to forecast CAT from gravity wave breaking, upper fronts and other atmospheric deformation regions. *Preprints, 6th Conference on Aviation Weather Systems*, Dallas, TX, American Meteorological Society, 509-514.

Marroquin, A., 1998: An advanced algorithm to diagnose atmospheric turbulence using numerical model output. *Preprints, 16th Conference on Weather Analysis and Forecasting*, Phoenix, AZ, American Meteorological Society.

McCann, D. W., 1997: A "novel" approach to turbulence forecasting. Preprints, *Seventh Conf. On Aviation, Range and Aerospace Meteorology*, Long Beach, CA, American Meteorological Society, 158-163.

Sharman, R, C. Tebaldi, and B. Brown, 1999: An integrated approach to clear-air turbulence forecasting. Preprints, *Eighth Conf. On Aviation, Range, and Aerospace Meteorology*, Dallas, TX, American Meteorological Society, 68-71.

Sharman, R., B. Brown, and S. Dettling, 2000: Preliminary results of the NCAR Integrated Turbulence Forecasting Algorithm (ITFA) to forecast CAT. *Preprints, 9th Conference on Aviation, Range, and Aerospace Meteorology*, 11-15 September, Orlando, FL, American Meteorological Society (Boston), 460-465.

APPENDIX A

AWC Turbulence Algorithm Evaluation – Winter 2001

Please fill out one questionnaire every day, considering turbulence at 18,000 ft and above.
Problems or questions: Contact Jennifer Mahoney at mahoney@fsl.noaa.gov (303-497-6514) or Barbara Brown at bgb@ucar.edu (303-497-8468).
NOTE: All displays are on the web at http://www-ad.fsl.noaa.gov/afra/rtvs/turb_2001/index.html.

Name:____________________________ Current Date (UTC):_________________

Current Time (UTC):___________

Date of turbulence (UTC) if different from above: _____________________

I. Weather Classification

Note: Using only PIREPs to identify turbulence events or turbulence outbreaks above 18,000 ft, please answer the following questions:

1. **How would you categorize this day, in terms of the extent of turbulence activity? (We are most interested in turbulence outbreaks and events during the hours 1200 to 0300 UTC.)**

   ___ A "big" day ___A "moderate" day ___A "small" day

   Comments:

2. **What was the maximum severity of turbulence today?** ___Chop ___Light ___Mod ___Severe

   Comments:

3. **What is the major cause of the turbulence?**
   ___Jet stream ___Mountain waves ___Convection
   ___Upper level trough ___Upper level ridge Other:__________________

4. **Was the turbulence located**
   ___ in a clear region? ___near-cloud?
   ___in-cloud? ___ can't determine

5. **Where is/was the turbulence located?** ***(Draw on map, if desired)***

___Over entire U.S. ___East Other:_________
___Northern half of U.S. ___Central ____________
___Southern half of U.S. ___West

6. **What were the most common altitude ranges for observed turbulence?**

____18-22,000 ft ___ 22-26,000 ft ___26-30,000 ft

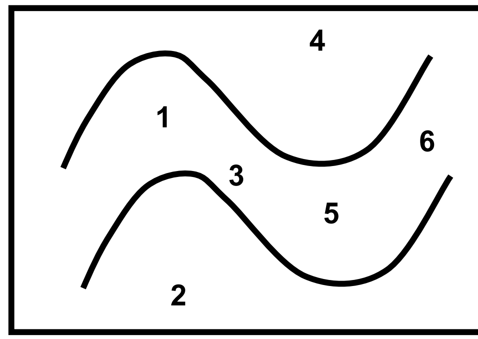
___30-34,000 ft ___34-38,000 ft ___38-42,000 ft

Comments:

7. **At what time of day did the major events occur?**
___12-15 UTC ___15-18 UTC ___18-21 UTC ___21-00 UTC
___Other:__________

8. **How long did the turbulence outbreaks persist (check all that apply)?**
___<1 hr ___1-2 hr ___2-3 hr ___4-6 hr ___6-12 hr ___>12 hr
___NA

9. **If turbulence related to troughs/ridges occurred, please use the graph to the right to identify the portion of the trough/ridge that was involved:**
___1 ___2 ___3 ___4 ___5 ___6



The chart I used to answer this questions is the

___250 mb chart ___300 mb chart

II. Algorithm Assessment

Still considering turbulence at levels of 18,000 ft and above, please review the algorithm output provided on the web site. Check all responses that apply.

1. **Which forecast issue and lead times did you consider?**
___12Z, 0-h ___12Z, 3-h ___12Z, 6-h ___12Z, 9-h ___12Z, 12-h
___15Z, 0-h ___15Z, 3-h ___15Z, 6-h ___15Z, 9-h
___18Z, 0-h ___18Z, 3-h ___18Z, 6-h
___21Z, 0-h ___21Z, 3-h

2. **Which algorithms …**

Index	***403***	***405***	***409***	***418***	***420***	***438***	***445***	***450***	***OTHER_______***
… did you consider?									
… captured turbulence well?									
… did *not* seem to capture the turbulence well?									
… seemed to forecast too much turbulence?									
… provided a good indication of turbulence associated with mountain waves?									

3. **The areal extent of the forecasts using a threshold of 0.5 (red and yellow areas), was, in general**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too small								
The correct size								
Too large								

4. **The forecasts generally indicated turbulence**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
North								
South								
East								
West								
Correct location								

5. **The forecasts (using a threshold of 0.5; red and yellow areas) indicated turbulence at altitudes that were**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too low								
Appropriate								
Too high								

6. **The turbulence severity indicated by the forecasts was, in general,**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too low								
Appropriate								
Too high								

7. **Do the problems with the forecasts seem to be related to biases in the numerical model or to an inaccurate forecast by the turbulence algorithms?**

___Numerical model   ___Algorithms   ___Can't determine

Comments:

8. **For the indices that captured the turbulence, please list any biases (e.g., varying regional capabilities) that you noticed.**

9. **Which numerical weather prediction model(s) were most helpful to you today for formulating any turbulence AIRMETs that you issued?**

10. **Any other comments (e.g., incorrect PIREPs)?**

APPENDIX B

Delta Turbulence Algorithm Evaluation – Winter 2001

Please fill out one questionnaire every day, considering turbulence at 18,000 ft and above.
Problems or questions: Contact Jennifer Mahoney at mahoney@fsl.noaa.gov (303-497-6514) or Barbara Brown at bgb@ucar.edu (303-497-8468).
NOTE: All displays are on the web at http://www-ad.fsl.noaa.gov/afra/rtvs/turb_2001/index.html.

Name:____________________________ Current Date (UTC):_________________ Current Time (UTC):___________

Date of turbulence (UTC) if different from above: ____________________

I. Weather Classification

Note: Using only PIREPs to identify turbulence events or turbulence outbreaks above 18,000 ft, please answer the following questions:

1. How would you categorize this day, in terms of the extent of turbulence activity? (We are most interested in turbulence outbreaks and events during the hours 1200 to 0300 UTC.)

___ A "big" day   ___A "moderate" day   ___A "small" day

Comments:

2. What was the maximum severity of turbulence today? ___Chop ___Light ___Mod ___Severe

Comments:

3. What is the major cause of the turbulence?

___Jet stream   ___Mountain waves   ___Convection
___Upper level trough   ___Upper level ridge   Other:__________________

4. Was the turbulence located

___ in a clear region?   ___near-cloud? ___in-cloud?
___ can't determine

5. Where is/was the turbulence located? ***(Draw on map, if desired)****.*

___Over entire U.S.   ___East   Other:________
___Northern half of U.S.   ___Central   ____________
___Southern half of U.S.   ___West

6. **What were the most common altitude ranges for observed turbulence?**

___18-22,000 ft ___ 22-26,000 ft ___26-30,000 ft

___30-34,000 ft ___34-38,000 ft ___38-42,000 ft

Comments:

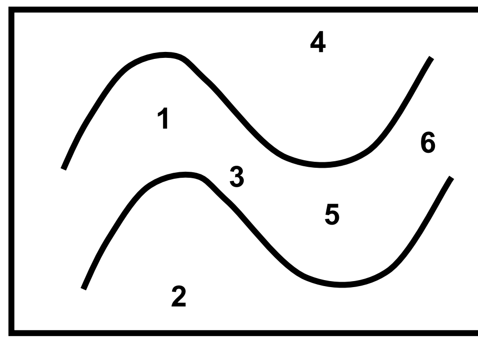
7. **At what time of day did the major events occur?**
___12-15 UTC ___15-18 UTC ___18-21 UTC ___21-00 UTC Other:__________

8. **How long did the turbulence outbreaks persist (check all that apply)?**
___<1 hr ___1-2 hr ___2-3 hr ___4-6 hr ___6-12 hr ___>12 hr
___NA

9. **If turbulence related to troughs/ridges occurred, please use the graph to the right to identify the portion of the trough/ridge that was involved:**
___1 ___2 ___3 ___4 ___5 ___6

The chart I used to answer this questions is the

___250 mb chart ___300 mb chart



II. Algorithm Assessment

Still considering turbulence at levels of 18,000 ft and above, please review the algorithm output provided on the web site. Check all responses that apply.

1. **Which forecast issue and lead times did you consider?**
___12Z, 0-h ___12Z, 3-h ___12Z, 6-h ___12Z, 9-h ___12Z, 12-h
___15Z, 0-h ___15Z, 3-h ___15Z, 6-h ___15Z, 9-h
___18Z, 0-h ___18Z, 3-h ___18Z, 6-h
___21Z, 0-h ___21Z, 3-h

2. **Which algorithms …**

Index	***403***	***405***	***409***	***418***	***420***	***438***	***445***	***450***	***OTHER_______***
… did you consider?									
… captured turbulence well?									
… did *not* seem to capture the turbulence well?									
… seemed to forecast too much turbulence?									
… provided a good indication of turbulence associated with mountain waves?									

3. **The areal extent of the forecasts using a threshold of 0.5 (red and yellow areas), was, in general**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too small								
The correct size								
Too large								

4. **The forecasts generally indicated turbulence**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
North								
South								
East								
West								
Correct location								

5. **The forecasts (using a threshold of 0.5; red and yellow areas) indicated turbulence at altitudes that were**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too low								
Appropriate								
Too high								

6. **The turbulence severity indicated by the forecasts was, in general,**

Index	*403*	*405*	*409*	*418*	*420*	*438*	*445*	*450*
Too low								
Appropriate								
Too high								

7. **Do the problems with the forecasts seem to be related to biases in the numerical model or to an inaccurate forecast by the turbulence algorithms?**
   ___Numerical model ___Algorithms ___Can't determine
   Comments:

8. **For the indices that captured the turbulence, please list any biases (e.g., varying regional capabilities) that you noticed.**

9. **Which numerical weather prediction model(s) were most helpful to you today for formulating any turbulence MetAlerts that you issued?**

10. **Any other comments (e.g., incorrect PIREPs)?**

Appendix C

AWC Forecaster Comments

Sheet #1

Wx Question #3: Near edge of cirrus from convection.
Algo Question #8: Most did well though 420, 438, and 445 all forecasted turbulence in KY and SWD that did not verify. 405, 409, 418, and 450 all did well by forecasting very little turbulence.

Sheet #3

Wx Question #1: Very few reports. 1 or 2 reports in southern CA. No real pattern or organization.
Algo Question #8: 409, 418, and 438 Forecast little turbulence over western US. All zeroed in on mid-Atlantic states where I forecast area from 180-290 only 1 or 2 PIREPs.

Sheet #4

Wx Question #1: No turbulence reports during time period (2100Z-0000Z)
Wx Question #2: No turbulence.
Algo Question #9: ETA, RUC2
Algo Question #10: 403, 409, and 418 showed small area Eastern KY and no other turbulence. 420 way over forecasted turbulence.

Sheet #6

Wx Question #1: No turbulence.
Algo Question #10: No turbulence. 409 and 438 did best since they forecasted little turbulence.

Sheet #7

Wx Question #1: No turbulence. Reported above light.
Algo Question #9: None issued.

Sheet #8

Algo Question #9: ETA and RUC2
Algo Question #10: Quiet day only 3 reports. 418 did best job.

Sheet #10

Wx Question #1: Very few PIREPs (only 2). One in RS one in TX.
Algo Question # 9: ETA
Algo Question #10: 403 captured NR planes turbulence well. 405 also good; it forecasted very little turbulence. 420 Way over forecast in terms of area affected. 438 did not over forecast as well.

Sheet #11

Algo Question #9: None issued

Sheet #13

Wx Question #2: Isolated report…very few moderate.
Algo Question #9: ETA and RUC.

Sheet #15

Wx Question #3: Deformation zone NW of trough.
Algo Question #8: 405 and 409 did best, even the over-forecasted altitudes.

Sheet #16

Wx Question #3:"Neck" at closed low.
Algo Question #8: 405 and 418 best job on altitude location.
Algo Question #9: ETA

Sheet #17
Algo Question #8: Most algorithms did well but some forecasted too much turbulence east and south of the main area.

Sheet #18

Algo Question #8: Of those that did well most also had a large, unverified area, south of the main area of turbulence reports.

Sheet #20

Wx Question #6: Only 1 report.

Sheet #21

Wx Question #1: 1 report
Algo Question #8: 405, 409, 418, and 450 all did well by forecasting very little turbulence. The others all had lots to the south and some had it north and west too.

Sheet #23

Algo Question #9: 403 would have bee useful if I had been using AIRMETs.

Sheet #24

Algo Question #8: Most did well with the areas of turbulence. However most also indicated a large area to the north. (over MN, ND and SD).

Sheet #25

Wx Question #3: Confluent flows

Sheet #26

Wx Question #1: Strong northwesterly low level winds and strong upper level trough across CO and northern NM. Several mountain waves reported
Wx Question #2: SIGMET x-ray was issued for occasional severe turbulence 270-410 southern CO and northern NM.
Algo Question #8: 403 and 445 very similar and did a good job without over forecasting.

Algo Question #9: ETA and RUC

Sheet #27

Wx Question #3: Confluent flow.
Algo Question #8: Most forecast turbulence north of area where winds were stronger rather than the confluent region.
Algo Question #9: RUC2
Algo Question #10: Confluent flow pattern and water vapor satellite big help.

Sheet #30

Wx Question #1: 9 reports moderate and 1 report severe.

Sheet #31

Algo Question #8: Bias to the south (405). Bias to the north (409). Bias to the west (445).

Sheet #32

Algo Question #9: RUC and ETA

Sheet #33

Algo Question #10: 450 forecasted very little turbulence.

Sheet #34

Algo Question #8: 418 least turbulence and closer to level.

Sheet #35

Algo Question #9: ETA and RUC.

Sheet #36

Algo Question #10: Only three reports of turbulence. Two in KY and one in IA.

Sheet #37

Wx Question #3: Shear zone north of polar jet.

Sheet #39

Wx Question #5: Evaluated central area.
Wx Question #6: No high level turbulence (21-00Z).
Algo Question #10: 403, 409, 420, and 445 all forecasted an area of turbulence over MS/AL area, which did not verify. The other algorithms had little or no turbulence indicated in that area. Since the turbulence forecast was in the eastern part of the area I evaluated. I listed the turbulence as too far east in item 4. This may not be the right way to list it. Same idea for 5 and 6.

Sheet #40

Algo Question #8: Most did well in capturing the main area in TX, but spread it eastward much too far. Some also highlighted an area in the Great Lakes where it was apparently smooth.

Sheet #41

Wx Question #4: Both in and out of cloud.
Algo Question #4: Very little turbulence forecast by 450.
Algo Question #9: 405 and 450 forecasted very little turbulence. These algorithms were the worst.

Sheet #43

Algo Question #8: Index 405 was the best today. Intensity and area coverage was about right. Index 418 was a close second but over forecasted.
Algo Question #9: RUC and ETA.

Sheet #44

Algo Question #8: 445 did best in terms of aerial coverage, intensity and altitudes.
Algo Question #9: ETA and RUC.

Sheet #45

Wx Question #3: Deformation zone.
Wx Question #5: Northwestern U.S..
Algo Question #8: 418 seemed to do the best in terms of area and altitudes.
Algo Question #9: ETA and RUC.

Sheet #46

Wx Question #5: Northeast.
Algo Question #8: 405-Bad-None forecast in Northeast U.S. 409 not much better in New England.

Sheet #47

Wx Question #2: Three chop reports and one turbulence report.
Wx Question #6: One report of turbulence. Two reports of chop.
Algo Question #10: Since there was only one report it was hard to answer some questions, e.g. when algorithms 403, 409, 420 and 445 forecasted a lot of turbulence over western Great Lakes and none occurred, it is hard to answer #5 since any altitude will be wrong-both too high and too low. 405 and 418 were best and 438 did pretty well. The others forecasted way too much turbulence over the western Great Lakes.

Sheet #49

Wx Question #2: Only one report within ±3 hours of 00Z.
Wx Question #5: Central AL.
Algo Question #7: All forecasted a swath of turbulence from TX panhandle to Southeast TX. None was reported.
Algo Question #8: Most were too intense for the area that had turbulence.

Algo Question #10: I checked multiple boxes in 2, 3, 4, 5 and 6 because, for the most part, the algorithms did well with the lone area of reported turbulence (though most were too intense). However the all forecasted a sizeable area of severe turbulence where none was reported. So in one sense the algorithms forecasted appropriate, but at the same time the areas were too large and since no turbulence was reported over TX presumably too intense also.

Sheet #50

Algo Question #10:18Z 290 in Southeast lower MI in error. CLE CWSU said it never got report and had no other reports.

Sheet #51

Wx Question #7: Maybe early too – don't know.
Algo Question #8: Forecasted turbulence over the northern Great Lakes. It was not verified.
Algo Question #10: I made two selections in a column several times. Often an algorithm captures the reported turbulence and also forecasts in area where no turbulence was reported. In such a case the algorithm did well and not so well. Often, an algorithm captures existing turbulence area well, but also forecasted another area which is of course too far North, South, East and West.

Sheet #52:

Wx Question #7: Continuing event from the first.
Algo Question #10: Crashed again.

Sheet #53

Algo Question #7: Most all of the algorithms forecasted main area of turbulence to be in Great Lakes region, where very little turbulence occurred. Most did a pretty good job indicating the area where turbulence was most prevalent.
Algo Question #8: Most that had turbulence in TX had it too far West. See above comments. Almost all had area too far to the North.
Algo Question #9: This evaluation was done on mtd shift, looking back at earlier data.

Sheet #54

Algo Question #8: 403 had too much turbulence in northern New England.

Algo Question #9: I worked high level but algorithm 409 looks like it did the best job today.

Sheet #55

Algo Question #9: ETA and RUC2.
Algo Question #10: 438 had appropriate size and area, but was too rare and too weak. Good on altitudes.

Sheet #56

Wx Question #6: Mostly in 20-30,000 feet.

Algo Question #8: Most too large.
Algo Question #9: RUC2 and ETA.
Algo Question #10: Good PIREP coverage. PC crashed once during assessment but was able to get back in.

Sheet #57

Wx Question #1: In the top 5-10% of turbulence days, in the winter, in central FL area.
Wx Question #2: SIGMETs for severe turbulence in all 3 FL areas, widespread.
Wx Question #6: Mostly between 24-33,000 feet.
Wx Question #7: Occurred over long time span, all day, but I only looked at 00Z ±3 hours.
Algo Question #7: Over forecasted area (SW TX) is an area with little traffic. Maybe it would have verified if there had been more reports.
Algo Question #8: Several (418 and 420) got the main area but also included area over the Great Lakes where no turbulence was reported.
Algo Question #9: Didn't work during this period-on midnight shift.
Algo Question #10: 405 and 418 were weakest at 30-34K feet, right where turbulence was very strong. Both 405 and 418 showed appropriate sized red areas in the layers above and below 30-34K feet. 420 depicted the Great Lakes being as turbulent as KS, MO and CO and there was nothing reported in the Great Lakes. Most algorithms forecasted turbulence too high, it appeared. However it may be that due to the severity of the turbulence no one wanted to climb through the turbulence and thus no reports in the higher layers.

Sheet #58:
Algo Question #9: RUC2.
Algo Question #10: Crashed PC at 03Z. Could not get back in.

Sheet #59

Wx Question #1: Big outbreak. Many moderate-severe or sever reports in central Florida between 21 -00Z. Even more in CO.
Wx Question #9: Similar to #1 but much less amplitude.
Algo Question #9: ETA and RUC2.
Algo Question #10: A big day. 405 too far south. 409 forecast turbulence far south and second area too far north. Would like to have looked at more algorithms but browser kept crashing. Could not look at 418 without browser crashing.

Sheet #60

Wx Question #9: Mostly southwesterly winds with trough to the northeast.
Algo Question #10: Could no get back into RTVS.

Sheet #61

Wx Question #1: SIGMET NE IA, MO and IL issued at 23Z.
Wx Question #3: Confluent split flow pattern.
Algo Question #9: RUC2 and ETA.

Algo Question #10: 405-21Z very good detection of turbulence area. Not so for 15Z. 420 caught severe turbulence area but also forecasted severe turbulence over a large area. 438 Very good as well. 405-21Z model run did well 15 and 18Z runs to small and too far southwest.

Sheet #62

Wx Question #1: Most turbulence near CNUTN.
Wx Question #2: Turbulence near Thunderstorm.
Algo Question #9: RUC2
Algo Question #10: 409 did okay at 15Z but did poorly at 18 and 21Z. 445 did best western CNUTN but was at altitudes unreal were about the reports. Most algorithms did not show any turbulence in area 2 (CNUTN). Unreal for our purposes bay not be bad since we forecaster for CNUTN turbulence.

Sheet #63

Algo Question #9: RUC2

Sheet #65

Wx Question #4: Cirrus clouds near area.
Wx Question #9: Not related to trough. In fast confluence flow.
Algo Question #9: ETA
Algo Question #10: 420 very poor turbulence everywhere. 438 very good best of all algorithms. Most algorithms too far west from turbulence forecasts.

Sheet #66

Algo Question #9: None – forecast for after 00Z.

Sheet #67

Algo Question #8: All seem to get some of the turbulence this was a “big” day.
Algo Question #10: 420 was the worst. 403 and 445 did the best. Still too much turbulence west side of trough and too much northwest of area.

Sheet #68

Algo Question #7: All missed event. It was a small area not a large outbreak.
Algo Question #8: Turbulence north of thunderstorm interacting with ridge. None did well.
Algo Question #9: Satellite only good data source.
Algo Question #10: 420 and 445 very bad. Way too much turbulence.

Sheet #69

Algo Question #8: 405 had very little turbulence forecast for the west. 438 very little turb forecast above FL260.

Sheet #70

Algo Question #8: 405 best area and only small amount of turbulence at correct altitude.

Sheet #71

Algo Question #8: 420 and 445 worst. 405 best (least turbulence)
Algo Question #10: Best forecast no turbulence (non-convective)

Sheet #72

Algo Question #10: Data not available on web page.

Sheet #73

Algo Question #10: Not applicable today.

Sheet #74

Wx Question #6: Only one report at FL370.
Algo Question #10: Not available. The FSL site was available but no turbulence algorithm outputs were available.

Sheet #75

Algo Question #8: None did well for this small area.
Algo Question #10: Over forecasted mostly north of area.

Sheet #76

Wx Question #6: Most turbulence 310-370.
Algo Question #8: None.
Algo Question #9: RUC2
Algo Question #10: 403 and 415 performed best. 420 over forecasted especially in the Great Lakes area.

Sheet #77

Wx Question #6: Mostly 20-37,000 ft.

Wx Question #7: All day.
Algo Question #8: 445 some too far north but not too bad. 409 best overall. 403 better at lower levels, worse higher up.
Algo Question #9: ETA
Algo Question #10: Good PIREP coverage all day starting at 12Z.

Sheet #78

Wx Question #6: Turbulence from 20-36,000 ft.
Algo Question #10: 438 was real bad.

Sheet #79

Wx Question #6: Light 24-39,000.
Algo Question #8: Over forecast by 405.
Algo Question #9: ETA

Sheet #80

Algo Question #10: 450 very little turbulence.